

Exponential-Time Approximation of Hard Problems

Marek Cygan, Łukasz Kowalik, Marcin Pilipczuk and Mateusz Wykurz*

Abstract

We study optimization problems that are neither approximable in polynomial time (at least with a constant factor) nor fixed parameter tractable, under widely believed complexity assumptions. Specifically, we focus on MAXIMUM INDEPENDENT SET, VERTEX COLORING, SET COVER, and BANDWIDTH.

In recent years, many researchers design exact exponential-time algorithms for these and other hard problems. The goal is getting the time complexity still of order $O(c^n)$, but with the constant c as small as possible. In this work we extend this line of research and we investigate whether the constant c can be made even smaller when one allows constant factor approximation. In fact, we describe a kind of approximation schemes — trade-offs between approximation factor and the time complexity.

We study two natural approaches. The first approach consists of designing a backtracking algorithm with a small search tree. We present one result of that kind: a $(4r - 1)$ -approximation of BANDWIDTH in time $O^*(2^{n/r})$, for any positive integer r .

The second approach uses general transformations from exponential-time exact algorithms to approximations that are faster but still exponential-time. For example, we show that for any reduction rate r , one can transform any $O^*(c^n)$ -time¹ algorithm for SET COVER into a $(1 + \ln r)$ -approximation algorithm running in time $O^*(c^{n/r})$. We believe that results of that kind extend the applicability of exact algorithms for NP-hard problems.

Classification: Algorithms and data structures; “fast” exponential-time algorithms

1 Introduction

Motivation One way of coping with NP-hardness is polynomial-time approximation, i.e. looking for solutions that are relatively close to optimal. Unfortunately it turns out that there are still many problems which do not allow for good approximation. Let us recall some examples. Håstad [16] showed that INDEPENDENT SET cannot be approximated in polynomial time with factor $n^{1-\epsilon}$ for any $\epsilon > 0$ unless $\text{NP} = \text{ZPP}$. The same holds for VERTEX COLORING due to Feige and Kilian [11]. By another result of Feige [8], SET COVER cannot be approximated in polynomial time with factor $(1 - \epsilon) \ln n$, where n is the size of the set to cover, for any $\epsilon > 0$ unless $\text{NP} \subseteq \text{DTIME}(n^{\log \log n})$.

Another approach is the area of parametrized complexity (see e.g. [7]). Then the goal is to find an algorithm with time exponential only in a parameter unrelated to the instance size (then we say the problem is *fixed parameter tractable*, FPT in short). This parameter may reflect complexity of the instance – like treewidth, but then we get an efficient algorithm only for some subclass of possible instances. Another choice of the parameter is the measure of the solution quality. For example, one can verify whether in an n -vertex graph there is a vertex cover of size k in $O(1.2738^k + kn)$ time [4]. Again, the parametrized approach does not succeed in some cases. Verifying whether a graph is k -colorable is NP-complete for any $k \geq 3$, while INDEPENDENT SET and SET COVER are $W[1]$ - and $W[2]$ -complete

*Institute of Informatics, University of Warsaw, Poland. This research is partially supported by a grant from the Polish Ministry of Science and Higher Education, project N206 005 32/0807. E-mail addresses: cygan@mimuw.edu.pl, kowalik@mimuw.edu.pl, malcin@mimuw.edu.pl, mateusz.wykurz@students.mimuw.edu.pl.

¹ $O^*(f(n))$ notation suppresses polynomial factors

respectively, meaning roughly that an FPT algorithm for INDEPENDENT SET or SET COVER would imply algorithms of that kind for a host of other hard problems.

The aforementioned hardness results motivate the study of “moderately exponential time” algorithms. The goal here is to devise algorithms with exponential running time $O(2^{n/r})$ with r big enough. Indeed, a $O(2^{n/50})$ -time algorithm may appear practical for some range of n , say $n \leq 1000$. Despite some progress in this area we are still far from exact algorithms with time complexity of that order. One of the most researched problems in this field is INDEPENDENT SET. Exhaustive search for that problem gives $O(2^n)$ time bound while the currently best published result [13] is $O(2^{n/3.47})$. For VERTEX COLORING the first $O^*(2^{n/0.77})$ -time algorithm by Lawler was then improved in a series of papers culminating in a breakthrough $O^*(2^n)$ bound of Björklund, Husfeldt and Koivisto [2].

Now consider the UNWEIGHTED SET COVER problem. The instance consists of a family of sets $\mathcal{S} = \{S_1, \dots, S_m\}$. The set $U = \bigcup \mathcal{S}$ is called *the universe* and we denote $n = |U|$. The goal is to find the smallest possible subfamily $\mathcal{C} \subseteq \mathcal{S}$ such that $\bigcup \mathcal{C} = U$. Assume that m is relatively small but big enough that finding an optimal solution using an exact algorithm is out of question, say $m = 150$. If the universe is small we can get a good approximation by the greedy algorithm (see e.g. [23]) with approximation ratio $H_n < \ln n + 1$. However this approximation guarantee becomes bad when n is big. A natural thing to consider is an approximation algorithm with better (e.g. constant) guarantee and with running time exponential but substantially lower than the best known exact algorithm. In this paper we explore such approach. Ideally, one would like to have a kind of trade-off between the running-time and approximation ratio – then one gets as much accuracy as can be afforded. We study two approaches yielding results of that kind.

Search Tree Techniques Many exponential-time algorithms (e.g. backtracking algorithms) can be viewed as visiting the nodes of an exponential sized search tree. The nodes of the tree correspond to instances of the problem, and typically the instances in the leaves are either trivial or at least solvable in polynomial-time.

A natural idea is to use a search tree which has fewer nodes than the search tree of the exact algorithm, and with leaves corresponding to instances that can be *approximated* in polynomial time. This natural approach was used in a work on MAX SAT by Dantsin, Gavrilovich, Hirsch and Konev [6].

In this paper we describe one result of that kind: a $(4r - 1)$ -approximation of BANDWIDTH in time $O^*(2^{n/r})$, for any positive integer r (see Section 2 for the definition of the problem and a brief discussion of known results).

This approach can be used also for INDEPENDENT SET and SET COVER. For example, INDEPENDENT SET has constant ratio approximation for graphs of bounded degree. A standard approach to exact algorithm for this problem (used e.g. in [13]) is a recursive algorithm, which picks a vertex v and checks two possibilities: (1) v is in the independent set – then removes v and its neighbors and makes a recursive call, and (2) v is not in the independent set – then removes v and makes a recursive call. When we can assume that the vertex picked is of large degree (if there is no such vertex polynomial-time approximation is used) we always get a big reduction of the instance size in one of the recursive calls which results in a better time bound than the exact algorithm. However, we will not elaborate on this because we obtained much better results (also for SET COVER) using the approach of reduction (see the next paragraph).

Reductions Consider an instance $\mathcal{S} = \{S_1, \dots, S_m\}$ of the UNWEIGHTED SET COVER problem described above. Assume m is even. Then create a new instance $\mathcal{Z} = \{Z_1, \dots, Z_{m/2}\}$ where $Z_i = S_{2i-1} \cup S_{2i}$. Next find an optimal solution $\text{OPT}_{\mathcal{Z}}$ for \mathcal{Z} using an exact algorithm. Let $\mathcal{C} = \{S_{2i-1} \mid Z_i \in \text{OPT}_{\mathcal{Z}}\} \cup \{S_{2i} \mid Z_i \in \text{OPT}_{\mathcal{Z}}\}$. Clearly, $|\text{OPT}_{\mathcal{Z}}| \leq |\text{OPT}_{\mathcal{S}}|$ and hence $|\mathcal{C}| \leq 2|\text{OPT}_{\mathcal{S}}|$. Thus we get a 2-approximation in $T(m/2)$ time where $T(m)$ is the best known bound for an exact algorithm, so currently just $T(m/2) = O^*(2^{m/2})$ after applying the exhaustive search. Of course this method is scalable – similarly we get a 5-approximation in time $O^*(2^{m/5})$. With the present computing speed it should allow to process instances with roughly 150 sets, even when the universe U is large.

The above phenomenon is the second of the two approaches studied this paper. We will call it a reduction. Let us state a more precise definition now (however, we will make it a bit more general

further in the paper). Consider a minimization problem P (reduction of a maximization problem is defined analogously) with a measure of instance size $s(I)$, and a measure of the solution quality $m(S)$. Let OPT_I be an optimal solution for instance I .

Let $r, a > 1$ be constants. A (r, a) -reduction (or simply a *reduction*) of problem P is a pair of algorithms called *reducer* and *merger* satisfying the following properties:

- Reducer transforms I , an instance of P , into a set $\{I_1, \dots, I_k\}$ of instances of P so that for every $i = 1, \dots, k$, $s(I_i) \leq s(I)/r + O(1)$.
- Let S_1, \dots, S_k be optimal solutions of instances I_1, \dots, I_k . Then Merger transforms the solutions S_i into S , a solution for instance I , so that $m(S) \leq a \cdot \text{OPT}_I$. (Merger may also use I and any information computed by reducer).

The constants r and a in the above definition will be called the *rate* and the *approximation* of the reduction, respectively (we assume $a \geq 1$, even for a maximization problem – then a solution with quality $\geq \text{OPT}_I/a$ is returned). Observe that above we described a $(2, 2)$ -reduction of UNWEIGHTED SET COVER. Actually, we noted that it generalizes to (r, r) -reduction for any $r \in \mathbb{N}$. If there is a $(r, a(r))$ -reduction for r being arbitrarily big we talk about *reduction scheme* and the function $a(r)$ is the *approximation* of this scheme. (This definition is very flexible, however, most of our schemes imply a reduction for any $r \in \mathbb{N}$, and sometimes even for any $r \in \mathbb{Q}$).

Problem	Approximation	Range of r	Current time bound
UNWEIGHTED SET COVER	$1 + \ln r$	$r \in \mathbb{Q}, r \geq 1$	$O^*(2^{n/r}), [2]$
SET COVER	$1 + \ln r$	$r \in \mathbb{Q}, r \geq 1$	$O^*(2^{n/(0.5r)} m^{\log n}),$ $O^*(2^{n/(0.31r)}), \text{App. E}$
SET COVER	r	$r \in \mathbb{N}, r \geq 1$	$O^*(2^{m/r}), [\text{folklore}]$
MIN DOMINATING SET	r	$r \in \mathbb{N}, r \geq 1$	$O^*(2^{0.598n/r}), [21]$
MAX INDEPENDENT SET	r	$r \in \mathbb{Q}, r \geq 1$	$O^*(2^{n/(3.47r)}), [13]$
COLORING ^a	$1 + \ln r$	$r \in \mathbb{Q} \cap [1, 4.05]$	$O^*(2^{n/(0.85r)}), [2]$
COLORING ^b	$1 + 0.247r \ln r$	$r \in \mathbb{Q}, r > 4.05$	$O^*(2^{n/(0.85r)}), [2]$
COLORING	r	$r \in \mathbb{N}, r > 1$	$O^*(2^{n/(0.85r)}), [2]$
BANDWIDTH	$r^{\log_2 9} = 9^k$	$r = 2^k, k \in \mathbb{N}$	$O^*(10^{n/r}), [10]$
SEMI-METRIC TSP	$1 + \log_2 r = 1 + k$	$r = 2^k, k \in \mathbb{N}$	$O^*(2^{n/(0.5r)}), [1]$

^aThis is an $O^*(2^{n/3.47})$ -time reduction by Björklund and Husfeldt, see [2]

^bThis is an $O^*(2^{n/(0.85r)})$ -time reduction.

Table 1: Our reductions. Last column shows time bounds of approximation algorithms obtained using the best known (polynomial space) exact algorithms.

We present reduction schemes for several most natural optimization problems that are both hard to approximate and resist FPT algorithms. Table 1 shows our results. In last column we put time bounds of approximation algorithms obtained by using our reductions with best known exact algorithms. As our motivations are partially practical, all these bounds refer to *polynomial space* exact algorithms (putting $r = 1$ gives the time complexity of the relevant exact algorithm). In most cases (BANDWIDTH, COLORING, INDEPENDENT SET, SEMI-METRIC TSP) there are faster exact exponential-space algorithms (hence we would get also faster approximations).

Note that by putting $r = n/\log n$ we get polynomial time approximations, and for SET COVER we get the approximation ratio $1 + \ln n$, which roughly matches the ratio of the (essentially optimal) greedy algorithm. Thus, our reduction can be viewed as a continuous scaling between the best possible polynomial time approximation and the best known exponential time algorithm. In other words, one can get as good solution as he can afford, by using as much time as he can. A similar phenomenon appears

in the case of SEMI-METRIC TSP. (This is not very surprising since these two reductions are based on the relevant polynomial time approximations).

The notion of reduction introduced in our paper is so natural that some reductions must have been considered before, especially for big reduction rates, like $r = n/\log n$, when the reductions essentially imply polynomial-time approximation algorithms. We are aware of one reduction described in the context of exponential-time approximation: Björklund and Husfeldt [2] described a reduction scheme with approximation $a(r) = (1 + \ln r)$ (worth using only for bounded values of r – see §C).

Related Work We have already mentioned results of Björklund and Husfeldt [2] and Dantsin et al. [6] on exponential-time approximation. The idea of joining the worlds of approximation algorithms and “moderately” exponential algorithms appeared also in a recent work of Vassilevska, Williams and Woo [22], however their direction of research is completely different from ours, i.e. they consider so-called hybrid algorithms. For example, they report to have an algorithm for BANDWIDTH which for given input *either* returns an $O(\log n)$ -approximate solution in polynomial time *or* returns a $(1 + \epsilon)$ -approximate solution in $O(2^{n/\log \log n})$ time. We see that the hybrid algorithm does not guarantee constant approximation ratio and hence cannot be directly compared with our work.

Another promising area is joining the worlds of parametrized complexity and polynomial-time approximation algorithms — see the survey paper [19].

Organization of the Paper We start from the approximation scheme for the BANDWIDTH problem in Section 2. Then in Section 3 we introduce slightly more general definition of reduction and then in Section 4 describe two reductions for SET COVER. (The reductions for MAXIMUM INDEPENDENT SET, VERTEX COLORING, BANDWIDTH and SEMI-METRIC TSP are put in Appendix due to space limitations.) We conclude in Section 5 by some complexity remarks that show relations between the notion of reduction and polynomial-time approximation.

2 Bandwidth

Let $G = (V, E)$ be an undirected graph. For a given *ordering of vertices*, i.e. a one-to-one function $f : V \rightarrow \{1, \dots, n\}$, its *bandwidth* is the maximum difference between the numbers assigned to the endpoints of an edge, i.e. $\max_{uv \in E} |f(u) - f(v)|$. *Bandwidth* of graph G , denoted by $\text{bw}(G)$, is the minimum possible bandwidth of an ordering. The BANDWIDTH problem asks to find, for a given graph, its bandwidth with the corresponding ordering.

BANDWIDTH is a notorious NP-hard problem. It was shown by Unger [20] that BANDWIDTH does not belong to APX even in very restricted case when G is a caterpillar, i.e. a very simple tree. It is also hard for any fixed level of the W hierarchy [3]. The best known polynomial-time approximation, due to Feige [9], has $O(\log^3 n \sqrt{\log n \log \log n})$ approximation guarantee. The fastest known exact algorithm works in time $O^*(5^n)$ and space $O^*(2^n)$ and it is due to Cygan and Pilipeczuk [5], while the best polynomial-space exact algorithm, due to Feige and Kilian [10], has time complexity $O^*(10^n)$.

We were able to find a reduction for the BANDWIDTH problem. Although it is probably the most nontrivial of our reductions, it gives, for any $k \in \mathbb{N}$, approximation ratio 9^k with reduction rate 2^k , which is far from being practical (see Appendix D for the details). As a corollary it gives a 9^k -approximation in time $O^*(10^{n/2^k})$ and polynomial space, or in time $O^*(5^{n/2^k})$ and $O^*(2^{n/2^k})$ space. It is an interesting open problem whether there is a better reduction for this problem.

Now we will describe a better approximation scheme using the approach of small search tree.

2.1 2-approximation in $O^*(3^n)$ -time (warm-up)

We begin with an algorithm which is very close to a fragment of the $O^*(10^n)$ -time exact algorithm of Feige and Kilian [10]. Assume w.l.o.g. the input graph is connected (we keep this assumption also in the following sections).

Let b be the bandwidth of the input graph — we may assume it is given, otherwise with just a $O(\log n)$ overhead using binary search one can find the smallest b for which the algorithm returns a solution.

Let us partition the set of positions $\{1, \dots, n\}$ into $\lceil n/b \rceil$ intervals of size b (except, possibly, for the last interval), so that for $j = 0, \dots, \lceil n/b \rceil - 1$, the j -th interval consists of positions $I_j = \{jb + 1, jb + 2, \dots, (j + 1)b\} \cap \{1, \dots, n\}$.

Pseudocode 2.1 Generating at most $n3^n$ assignments in the 2-approximation algorithm.

```

1: procedure GENERATEASSIGNMENTS( $A$ )
2:   if all nodes in are assigned then
3:     If each interval  $I_j$  is assigned  $|I_j|$  vertices, order the vertices in intervals arbitrarily and return
       the ordering.
4:   else
5:      $v \leftarrow$  a vertex with a neighbor  $w$  already assigned.
6:     if  $A(w) > 0$  then GENERATEASSIGNMENTS( $A \cup \{(v, A(w) - 1)\}$ )
7:     GENERATEASSIGNMENTS( $A \cup \{(v, A(w))\}$ )
8:     if  $A(w) < \lceil n/b \rceil - 1$  then GENERATEASSIGNMENTS( $A \cup \{(v, A(w) + 1)\}$ )

9: procedure MAIN
10:  for  $j \leftarrow 0$  to  $\lceil n/b \rceil - 1$  do
11:    GENERATEASSIGNMENTS ( $\{(r, j)\}$ )       $\triangleright$  Generate all assignments with  $r$  in interval  $I_j$ 

```

The algorithm finds a set of assignments of vertices into intervals I_j such that if there is an ordering π of bandwidth at most b , for at least one of these assignments, for every vertex v , $\pi(v)$ lies in the assigned interval. Clearly, if there is an ordering of bandwidth b , at least one such assignment exists. The following method (introduced by Feige and Kilian originally for intervals of length $b/2$) finds the required set with only at most $n3^n$ assignments. Choose an interval for the first vertex r (chosen arbitrarily) in all $\lceil n/b \rceil$ ways. Then pick vertices one by one, each time taking a vertex adjacent to an already assigned one. Then there are at most 3 intervals where the new vertex can be put and so on. (See Pseudocode 2.2. (Partial) assignment is represented by a set of pairs; a pair (v, j) means that $A(v) = I_j$.)

Obviously, when there are more vertices assigned to an interval than its length, the assignment is skipped. Now it is clear that for any remaining assignment *any* ordering of the vertices inside intervals gives an ordering of bandwidth at most $2b$. Hence we have a 3-approximation in $O^*(3^n)$ -time and polynomial space.

Note also that if we use intervals of length $b/2$ as in Feige and Kilian's algorithm, similar method gives 3/2-approximation in $O^*(5^n)$ -time.

2.2 Introducing the Framework

We are going to extend the idea from the preceding section further. To this end we need generalized versions of simple tools used in the 2-approximation above.

Let a (partial) interval assignment be any (partial) function $A : V \rightarrow 2^{\{1, \dots, n\}}$, that assigns intervals of positions to vertices of the input graph. An interval $\{i, i + 1, \dots, j\}$ will be denoted $[i, j]$. The size of an interval is simply the number of its elements. Let $\pi : V \rightarrow \{1, \dots, n\}$ be an ordering. When for every vertex v , $\pi(v) \in A(v)$, we will say that an interval assignment A is *consistent* with π and π is consistent with A .

In Section 2.1 all intervals had the same size b , moreover two intervals were always either equal or disjoint. When this latter condition holds, it is trivial to verify whether there is an assignment consistent a given interval assignment — it suffices to check whether the number of vertices assigned to any interval does not exceed its size. Luckily, in a general case it is still possible in polynomial time: just note that this is a special case of scheduling jobs on a single machine with release and deadline times specified

for each job (see e.g. [18], Sec. 4.2) — hence we can use the simple greedy algorithm which processes vertices in the order of $\max A(v)$ and assigns each vertex to the smallest available position.

Proposition 2.1. *For any interval assignment A one can verify in $O(n \log n)$ time whether there is an ordering consistent with A .*

To get a nice approximation ratio, however, we need a bound on the bandwidth of resulting ordering. The obvious bound is

$$\max_{uv \in E} \max_{\substack{i \in A(u), \\ j \in A(v)}} |i - j|.$$

The above bound was sufficient in Section 2.1, but we will need a better bound.

Lemma 2.2. *Let A be an interval assignment for an input graph $G = (V, E)$. Let s be the size of the largest interval in A , i.e. $s = \max_{v \in V} |A(v)|$. If there is an ordering π^* of bandwidth b consistent with A , then one can find in polynomial time an ordering π that is consistent with A and has bandwidth at most $s + b$.*

Proof. Consider any edge uv . Clearly, $\pi^*(u) \in [\min A(v) - b, \max A(v) + b]$. Hence we can replace $A(u)$ by $A(u) \cap [\min A(v) - b, \max A(v) + b]$, maintaining the invariant that π^* is consistent with A . Similarly, we can replace $A(v)$ by $A(v) \cap [\min A(u) - b, \max A(u) + b]$. Our algorithm performs such replacements for every edge $uv \in E$. As a result we get an assignment A' such that for every edge uv , $\max_{i \in A'(u), j \in A'(v)} |i - j| \leq s + b$. It is clear that any ordering consistent with A' has bandwidth at most $s + b$. Such the ordering can be found in polynomial time by Proposition 2.1 \square

In the following sections it will be convenient to formalize a little the order in which intervals are assigned to the vertices of the input graph. Recall that each time an interval is assigned to a new vertex v , v has an already assigned neighbor w , except for the initial vertex r . In other words, the algorithm builds a rooted spanning tree for each assignment (here, r is the root, and w is the parent of v). In what follows, we will fix a rooted spanning tree T , and our algorithm will generate interval assignments in such a way that the first vertex assigned r is the root of T , and whenever a vertex $v \neq r$ is assigned, its parent in T has been already assigned.

2.3 3-approximation in $O^*(2^n)$ -time

This time, the algorithm uses $\lceil n/b \rceil$ intervals, each of size $2b$ (except for one or two last intervals), so that for $j = 0, \dots, \lceil n/b \rceil - 1$, the j -th interval consists of positions $I_j = \{jb + 1, jb + 2, \dots, (j+2)b\} \cap \{1, \dots, n\}$. Note that the intervals overlap.

Pseudocode 2.2 Generating at most $n2^n$ assignments in the 3-approximation algorithm.

```

1: procedure GENERATEASSIGNMENTS( $A$ )
2:   if all nodes in  $T$  are assigned then
3:     Using Lemma 2.2 find ordering consistent with interval assignment corresponding to  $A$ .
4:   else
5:      $v \leftarrow$  a node in  $T$  such that  $v$ 's parent  $w$  is assigned.
6:     if  $A(w) > 0$  then GENERATEASSIGNMENTS( $A \cup \{(v, A(w) - 1)\}$ )
7:     if  $A(w) < \lceil n/b \rceil - 1$  then GENERATEASSIGNMENTS( $A \cup \{(v, A(w) + 1)\}$ )

8: procedure MAIN
9:   for  $j \leftarrow 0$  to  $\lceil n/b \rceil - 1$  do
10:    GENERATEASSIGNMENTS ( $\{(r, j)\}$ ) ▷ Generate all assignments with root in  $I_j$ 

```

The algorithm generates all possible assignments of vertices to intervals in such a way that if a node in T is assigned to interval I_j then each of its children is assigned to interval I_{j-1} or I_{j+1} . Clearly, there are at most $n2^n$ such assignments. Moreover, if there is an ordering π of bandwidth b , then the algorithm

generates an interval assignment A_π consistent with π . To find A_π , visit the nodes of T in preorder, assign the root r to the interval $S_{\lfloor (\pi(r)-1)/r \rfloor}$, and for each node v with parent w already assigned to an interval I_j , put $A_\pi(v) = I_{j+1}$ if $\pi(v) > (j+1)b$ and $A_\pi(v) = I_{j-1}$ otherwise. For each generated assignment the algorithm tries to find an ordering of bandwidth at most $3b$ using Lemma 2.2. Clearly, it succeeds for at least one assignment, namely A_π . The algorithm is sketched in Pseudocode 2.2.

2.4 $(4r-1)$ -approximation in $O^*(2^{n/r})$ -time

In this section we are going to generalize the algorithm from the prior section to an approximation scheme. Let r be a positive integer. We will describe a $(4r-1)$ -approximation algorithm. Our algorithm uses intervals of sizes $2ib$, for $r \leq i \leq 2r-1$. Note that unlike in previous algorithms, intervals of many different sizes are used. As before, intervals will begin in positions $jb+1$, for $j = 0, \dots, \lceil n/b \rceil - 1$. The interval beginning in $jb+1$ and of length $2ib$ will be denoted by $I_{j,2i}$. For convenience, we allow intervals not completely contained in $\{1, \dots, n\}$, but each assigned interval contains at least one position from $\{1, \dots, n\}$.

Pseudocode 2.3 Generating assignments in the $(4r-1)$ -approximation algorithm.

```

1: procedure GENERATEASSIGNMENTS( $A$ )
2:   if all nodes in  $T$  are assigned then
3:     Cut all intervals in  $A$  to make them contained in  $\{1, \dots, n\}$ .
4:     Using Lemma 2.2 find ordering consistent with interval assignment corresponding to  $A$ .
5:   else
6:      $v \leftarrow$  a node in  $T$  such that  $v$ 's parent  $w$  is assigned; let  $I_{j,2i} = A(w)$ .
7:     if  $i+1 \leq 2r-1$  then
8:       GENERATEASSIGNMENTS( $A \cup \{(v, I_{j-1,2(i+1)})\}$ )
9:     else
10:      if  $j-1+2r \geq 1$  then GENERATEASSIGNMENTS( $A \cup \{(v, I_{j-1,2r})\}$ )
11:      if  $j-1+2r \leq \lceil n/b \rceil - 1$  then GENERATEASSIGNMENTS( $A \cup \{(v, I_{j-1+2r,2r})\}$ )
12: procedure MAIN( $i_0$ )
13:   for  $j \leftarrow 0$  to  $\lceil n/b \rceil - 1$  do
14:     GENERATEASSIGNMENTS( $\{(r, I_{j,2i_0})\}$ )  $\triangleright$  Generate all assignments with root in  $I_{j,2i_0}$ 

```

The algorithm is sketched in Pseudocode 2.3. Let $i_0 \in \{r, \dots, 2r-1\}$ be a parameter that we will determine later. The algorithm assigns the root of T to all possible intervals of size $2i_0b$ that overlap with $\{1, \dots, n\}$ and extends each of these partial assignments recursively.

To extend a given assignment, the algorithm chooses a node v of T such that the parent w of v has been already assigned an interval, say $I_{j,2i}$. Consider the interval $I_{j-1,2(i+1)}$ which is obtained from $I_{j,2i}$ by “extending” it by b positions both at left and right side. Note that in any ordering consistent with the current assignment, v is put in a position from $I_{j-1,2(i+1)}$. Hence, if $I_{j-1,2(i+1)}$ is not too big, i.e. $i+1 \leq 2r-1$, the algorithm simply assigns $I_{j-1,2(i+1)}$ to v and proceeds with no branching (just one recursive call). Otherwise, if $i+1 = 2r$, the interval $I_{j-1,2(i+1)}$ is split into two intervals of size $2r$, namely $I_{j-1,2r}$ and $I_{j-1+2r,2r}$ and two recursive calls follow: with v assigned to $I_{j-1,2r}$ and $I_{j-1+2r,2r}$ respectively.

As before, for every generated assignment (after cutting the intervals to make them contained in $\{1, \dots, n\}$) the algorithm applies Lemma 2.2 to verify whether it is consistent with an ordering of bandwidth $[2(2r-1)+1]b = (4r-1)b$. Similarly as in the case $r=1$ described before, for at least one assignment such the ordering is found.

We conclude with the time complexity analysis. Observe that the nodes at tree distance d from the root are assigned intervals of size $2[(i_0+d) \bmod r + r]$. It follows that branching appears only when $i_0 + d \equiv 0 \pmod{r}$. Let $\hat{n}(i_0)$ denote the number of nodes at tree distance d satisfying this condition.

It is clear that the above algorithm works in time $O^*(2^{\hat{n}(i_0)})$. Since $\sum_{i \in \{r, \dots, 2r-1\}} \hat{n}(i) = n$, for some $i \in \{r, \dots, 2r-1\}$, $\hat{n}(i) \leq n/r$. By choosing this value as i_0 , we get the $O^*(2^{n/r})$ time bound.

Theorem 2.3. *For any positive integer r , there is a $(4r-1)$ -approximation algorithm for BANDWIDTH running in $O^*(2^{n/r})$ time and polynomial space. \square*

3 Reducibility (slightly more general)

In this section we introduce a slightly more general version of reduction and discuss some of its basic properties. Essentially the difference from the version in the Introduction is that sometimes between reducer and merger we want to use approximation algorithm instead of exact one.

As before let P be a minimization problem, let $s(I)$ and $m(S)$ denote the measures of the instance size and the solution quality, respectively. Let $r > 1$ be a constant and $f : \mathbb{R} \rightarrow \mathbb{R}$ a function. An (r, f) -reduction (or simply a *reduction*) of P is a pair of algorithms called *reducer* and *merger* satisfying the following properties:

- Reducer transforms I , an instance of P , into a set $\{I_1, \dots, I_k\}$ of instances of P so that for every $j = 1, \dots, k$, $s(I_j) \leq s(I)/r + O(1)$.
- Let S_1, \dots, S_k be solutions of instances I_1, \dots, I_k . Let $\alpha > 1$ be an approximation guarantee of these solutions, i.e. for $j = 1, \dots, k$, $m(S_j) \leq \alpha m(\text{OPT}_{I_j})$. Then merger transforms the solutions S_i into S , a solution for instance I , so that $m(S) \leq f(\alpha) \text{OPT}_I$. (Merger may also use I and any information computed by reducer).

As before, r is called the rate and f is called the approximation (since we do not expect f to be a constant function it should not lead to ambiguity). Again, if there is a (r, f_r) -reduction for arbitrarily big r we deal with *reduction scheme* with approximation $a(r, \alpha) = f_r(\alpha)$. Note that the already described reduction for UNWEIGHTED SET COVER has approximation $a(r, \alpha) = r\alpha$.

The time complexity of the reduction is the sum of (worst-case) time complexities of reducer and merger. In most cases our reductions will be polynomial time. However, under some restrictions exponential-time reductions may be interesting as well.

The following lemma will be useful (an easy proof is in Appendix A):

Lemma 3.1 (Reduction Composition). *If there is an (r, f) -reduction R then for any positive $k \in \mathbb{N}$ there is a (r^k, f^k) -reduction² R' for the same problem. Moreover, if the merger of R generates a polynomial number of instances then R' has the same time complexity as R , up to a polynomial factor.*

Note that the Reduction Composition Lemma implies that once we find a single (r, f) -reduction, it extends to a reduction scheme, though for quite limited choice of r . We will see more consequences of this lemma in Section 5.

4 Set Cover

We will use the notation for SET COVER from the Introduction. Since here we consider the general, weighted version of the problem, now each set $S \in \mathcal{S}$ comes with its weight $w(S)$. We will also write $w(\mathcal{C})$ for the total weight of a family of sets \mathcal{C} . In the case of SET COVER there are two natural measures for size of the instance: the size of the universe U and the number of sets in the family \mathcal{S} . We will present reductions for both measures. Both reductions work for the weighted version of the problem.

² f^k is the composition: $f^k = f \circ f^{k-1}$, $f^0 = \text{id}$.

4.1 Reducing the size of universe

An r -approximate solution of SET COVER can be found by dividing U into r parts, covering each of them separately and returning the union of these covers. It corresponds to a reduction scheme with approximation $a(r, \alpha) = r\alpha$ for $r \in \mathbb{N}$. However, we will describe a much better reduction scheme.

Let's recall the greedy algorithm (see e.g. [23]), called **Greedy** from now. It selects sets to the cover one by one. Let \mathcal{C} be the family of sets chosen so far. Then **Greedy** takes a set that covers *new* elements as cheap as possible, i.e. chooses S so as to minimize $w(S)/|S \setminus \bigcup \mathcal{C}|$. For each element $e \in S \setminus \bigcup \mathcal{C}$ the amount $w(S)/|S \setminus \bigcup \mathcal{C}|$ is called the *price* of e and denoted as $\text{price}(e)$. This procedure continues until \mathcal{C} covers the whole U . Let e_1, \dots, e_n be the sequence of all elements of U in the order of covering by **Greedy** (ties broken arbitrarily). The standard analysis of **Greedy** uses the following lemma (see [23] for the proof).

Lemma 4.1. *For each $k \in 1, \dots, n$, $\text{price}(e_k) \leq w(\text{OPT})/(n - k + 1)$*

The idea of our reduction is very simple. For example, assume we aim at a reduction rate 2, and n is even. Lemma 4.1 tells us that **Greedy** starts from covering elements very cheaply, and then pays more and more. So we just stop it before it pays much but after it covers sufficiently many elements. Note that if we manage to stop it just after $e_{n/2}$ is covered the total price of the covered elements (and hence the weight of the sets chosen) is at most $(H_n - H_{n/2})w(\text{OPT}) = (\ln 2 + O(1/n))w(\text{OPT})$. If we cover the remaining elements, say, by exact algorithm we get roughly a $(1 + \ln 2)$ -approximation. However the set that covers $e_{n/2}$ may cover many elements e_i , $i > n/2$. By Lemma 4.1 the price of each of them is at most $w(\text{OPT})/(n/2) = 2w(\text{OPT})/n$. Hence this last set costs us at most $w(\text{OPT})$ and together we get roughly a $(2 + \ln 2)$ -approximation. Luckily, it turns out that paying $w(\text{OPT})$ for the last set chosen by **Greedy** is not necessary: below we show a refined algorithm which would yield a $(1 + \ln 2)$ -approximation in this particular case.

Theorem 4.2. *There is a polynomial-time $|U|$ -scaling reduction scheme for SET COVER with approximation $a(r, \alpha) = \alpha + \ln r + O(1/n)$, for any $r \in \mathbb{Q}$, $r > 1$.*

Pseudocode 4.1 Universe-scaling reducer for SET COVER

```

1:  $\mathcal{C} \leftarrow \emptyset$ .
2: while  $\bigcup \mathcal{S} \cup \bigcup \mathcal{C} = U$  do
3:   Find  $T \in \mathcal{S}$  so as to minimize  $\frac{w(T)}{|T \setminus \bigcup \mathcal{C}|}$ 
4:   if  $n - |\bigcup \mathcal{C} \cup T| > n/r$  then
5:      $\mathcal{C} \leftarrow \mathcal{C} \cup \{T\}$ .
6:   else
7:      $\mathcal{C}_T \leftarrow \mathcal{C}$ 
8:     [Create an instance  $I_T = (\mathcal{S}_T, w)$ :]
9:     for each  $P \in \mathcal{S}$ ,  $\mathcal{S}_T$  contains set  $P \setminus (\bigcup \mathcal{C}_T \cup T)$ , of weight  $w(P)$ .
10:     $\mathcal{S} \leftarrow \mathcal{S} \setminus \{T\}$ 

```

Proof. Let $I = (\mathcal{S}, w)$ be an instance of SET COVER problem. Reducer works similarly as **Greedy**. However, before adding a set T to the partial cover \mathcal{C} it checks whether adding T to \mathcal{C} makes the number of non-covered elements at most n/r . If so, T is called a *crossing set*. Instead of adding T to \mathcal{C} , reducer creates an instance $I_T = (\mathcal{S}_T, w)$ of SET COVER that will be used to cover the elements covered neither by \mathcal{C} nor by T . Namely, for each $P \in \mathcal{S}$, \mathcal{S}_T contains set $P \setminus (\bigcup \mathcal{C} \cup T)$, of weight $w(P)$. Apart from I_T reducer stores \mathcal{C}_T , a copy of \mathcal{C} , which will be used by merger. After creating the instance, set T is removed from the family of available sets \mathcal{S} . If it turns out that the universe cannot be covered after removing T , i.e. $\bigcup \mathcal{S} \cup \bigcup \mathcal{C} \neq U$, the reducer stops. See Pseudocode 4.1 for details. Note that reducer creates at least 1 and at most $|\mathcal{S}|$ instances.

Let I_T be any instance created for some crossing set T and let $\text{SOL}_T \subset \mathcal{S}$ be its solution such that $w(\text{SOL}_T) \leq \alpha w(\text{OPT}_{I_T})$, $\alpha \geq 1$. Let $\mathcal{S}'_T = \mathcal{C}_T \cup \{T\} \cup \text{SOL}_T$. Clearly, \mathcal{S}'_T is a cover of U for every crossing set T . The merger simply selects the lightest of these covers.

Let T^* be the first crossing set found by reducer such that T^* belongs to OPT_I , some optimal solution for instance I (note that at least one crossing set is in OPT_I). Clearly $\text{OPT}_I \setminus \{T^*\}$ covers $\bigcup \mathcal{S}_{T^*}$. Hence $w(\text{OPT}_{I_{T^*}}) \leq w(\text{OPT}_I \setminus \{T^*\})$ so $w(T^*) + w(\text{OPT}_{I_{T^*}}) \leq w(\text{OPT}_I)$. It follows that $w(T^*) + w(\text{SOL}_{T^*}) \leq \alpha w(\text{OPT}_I)$. Since \mathcal{C}_{T^*} covers less than $n - n/r$ elements, by Lemma 4.1

$$\begin{aligned} w(\mathcal{C}_{T^*}) &\leq \sum_{k=1}^{\lfloor n-n/r \rfloor} \frac{w(\text{OPT}_I)}{n-k+1} = \sum_{k=1}^{n-\lceil n/r \rceil} \frac{w(\text{OPT}_I)}{n-k+1} = (H_n - H_{\lceil n/r \rceil})w(\text{OPT}_I) = \\ &= (\ln n - \ln \lceil n/r \rceil + O(1/n))w(\text{OPT}_I) \leq (\ln r + O(1/n))w(\text{OPT}_I). \end{aligned}$$

We conclude that merger returns a cover of weight $\leq (\alpha + \ln r + O(\frac{1}{n}))w(\text{OPT}_I)$. \square

Clearly, to make use of universe-scaling reduction we need a $O^*(c^n)$ exact algorithm, where c is a constant. As far as we know there is no such result published. However we can follow the divide-and-conquer approach of Gurevich and Shelah [15] rediscovered recently Björklund and Husfeldt [1] and we get a $O^*(4^n m^{\log n})$ -time algorithm. If m is big we can use another, $O^*(9^n)$ -time version of it. See Appendix E for details.

We also note that for the unweighted case there is an $O(2^n mn)$ -time polynomial space algorithm by Björklund et al. [2] using the inclusion-exclusion principle.

4.2 Reducing the number of sets

Recall the reduction described in the introduction. In the weighted version it fails, basically because the sets from the optimal solution may be joined with some heavy sets. The natural thing to do is sorting the sets according to their weight and joining only neighboring sets. This simple modification does not succeed fully but with some more effort we can make it work.

Theorem 4.3. *There is a polynomial-time $|\mathcal{S}|$ -scaling reduction scheme for SET COVER with approximation $a(r, \alpha) = \alpha r$, for any $r \in \mathbb{N}$, $r > 1$.*

Proof. Reducer starts from sorting the sets in \mathcal{S} in the order of non-decreasing weight. So let $\mathcal{S} = \{S_1, \dots, S_m\}$ so that $w(S_1) \leq w(S_2) \leq \dots \leq w(S_m)$. Next it partitions this sequence into blocks \mathcal{B}_i , $i = 1, \dots, \lceil m/r \rceil$, each of size at most r , namely $\mathcal{B}_i = \{S_j \in \mathcal{S} \mid (i-1)r < j \leq ir\}$. Let $U_i = \bigcup \mathcal{B}_i$ be the union of all sets in \mathcal{B}_i and define its weight as the total weight of \mathcal{B}_i , i.e. $w(U_i) = w(\mathcal{B}_i)$. For any $k = 1, \dots, m$ we also define $\mathcal{X}_k = \{S_j \in \mathcal{B}_{\lceil k/r \rceil} \mid j < k\}$ and $V_k = \bigcup \mathcal{X}_k$ with $w(V_k) = w(\mathcal{X})$. Reducer creates m instances, namely $\mathcal{S}_i = \{U_j \mid S_i \notin \mathcal{B}_j\} \cup \{V_i, S_i\}$ for $i = 1, \dots, m$.

Of course any subfamily (or a cover) $\mathcal{C} \subseteq \mathcal{S}_i$ corresponds to $\widehat{\mathcal{C}}$, a subfamily of \mathcal{S} with the same weight obtained from \mathcal{C} by splitting the previously joined sets (we will use this denotation further). Clearly $\bigcup \mathcal{C} = \bigcup \widehat{\mathcal{C}}$, in particular if \mathcal{C} is a cover, so is $\widehat{\mathcal{C}}$.

Let $\mathcal{C}_1, \dots, \mathcal{C}_m$ be solutions (covers) for the instances created by reducer, such that $w(\mathcal{C}_i) \leq \alpha w(\text{OPT}_{\mathcal{S}_i})$. Merger simply chooses the lightest of them, say \mathcal{C}_q , and returns $\widehat{\mathcal{C}_q}$, a cover of U .

Now it suffices to show that one of the instances has a cover that is light enough. Let $i^* = \max\{i \mid S_i \in \text{OPT}\}$. We focus on instance \mathcal{S}_{i^*} . If $\mathcal{X}_{i^*} \cap \text{OPT} = \emptyset$ we choose its cover $\mathcal{R} = \{U_j \in \mathcal{S}_{i^*} \mid \mathcal{B}_j \cap \text{OPT} \neq \emptyset\} \cup \{S_{i^*}\}$, otherwise $\mathcal{R} = \{U_j \in \mathcal{S}_{i^*} \mid \mathcal{B}_j \cap \text{OPT} \neq \emptyset\} \cup \{V_{i^*}, S_{i^*}\}$. Clearly it suffices to show that $w(\widehat{\mathcal{R}} \setminus \text{OPT}) \leq (r-1)w(\text{OPT})$. Consider any $S_i \in \widehat{\mathcal{R}} \setminus \text{OPT}$. If $S_i \notin \mathcal{X}_{i^*}$ we put $f(i) = \min\{j \mid S_j \in \text{OPT} \text{ and } \lceil j/r \rceil > \lceil i/r \rceil\}$, otherwise $f(i) = i^*$. Then $w(S_i) \leq w(S_{f(i)})$. We see that f maps at most $r-1$ elements to a single index of a set from OPT , so indeed $w(\widehat{\mathcal{R}} \setminus \text{OPT}) \leq (r-1)w(\text{OPT})$ and hence $w(\widehat{\mathcal{R}}) \leq rw(\text{OPT})$. It follows that $w(\text{OPT}_{\mathcal{S}_{i^*}}) \leq rw(\text{OPT})$ so $w(\mathcal{C}_{i^*}) \leq \alpha rw(\text{OPT})$ and finally $w(\widehat{\mathcal{C}_q}) \leq \alpha rw(\text{OPT})$. \square

4.3 Special Case: (Weighted) Minimum Dominating Set

Of course, MINIMUM DOMINATING SET is a special case of SET COVER – a graph $G = (V, E)$ corresponds to the set system $\mathcal{S} = \{N[v] \mid v \in V\}$, where $N[v]$ consists of v and its neighbors. Note the set merging algorithm described in Section 4.2 can be adapted here: merging sets corresponds simply to identifying vertices. Hence we get a reduction scheme with approximation $a(r, \alpha) = \alpha r$. Combined with the recent $O(2^{0.598n})$ -time exact algorithm by Rooij and Bodlaender we get an r -approximation in time $O(2^{0.598n/r})$, for any natural r .

5 Reductions and polynomial-time approximation

Is it possible to improve any of the reductions presented before? Are some of them in some sense optimal? To address these questions at least partially we explore some connections between reductions and polynomial-time approximation. For example note that the $(r, \alpha r)$ -reduction for MAX INDEPENDENT SET implies $(n/\log n)$ -approximation in polynomial time, by putting $r = n/\log n$ and using an exact algorithm for instances of size $O(\log n)$. Since we know that MAX INDEPENDENT SET cannot be approximated much better in polynomial time it suggests that this reduction may be close to optimal in some sense. The following lemma is an immediate consequence of Reduction Composition Lemma. Let us call a reduction *bounded* when the reducer creates $O(1)$ instances.

Lemma 5.1. *If for some $r > 1$ there is a polynomial-time bounded (r, f) -reduction for problem P , then P is $f^{\log_r n - \log_r \log_2 n}(1)$ -approximable in polynomial time.*

Corollary 5.2. *If for some constants $c, r, r > 1, c > 0$ there is a polynomial-time bounded (r, f) -reduction with $f(\alpha) = c\alpha + o(\alpha)$ for problem P , then P is $O(\frac{n}{\log n})^{\log_r c}$ -approximable in polynomial time.*

Note that Corollary 5.2 implies that neither MAX INDEPENDENT SET nor VERTEX COLORING has a polynomial time bounded $(r, qr\alpha + o(\alpha))$ -reduction for any $q < 1$, unless $\text{NP} = \text{ZPP}$. If we skip the assumption that the reduction is bounded, the existence of such a reduction implies approximation in $n^{O(\log n)}$ time, which is also widely believed to be unlikely. Hence improved reductions need use either exponential time or some strange dependence on α , say $a(\alpha) = 0.1\alpha^2$.

References

- [1] A. Björklund and T. Husfeldt. Exact algorithms for exact satisfiability and number of perfect matchings. In *Proc. ICALP'06*, pages 548–559, 2006.
- [2] A. Björklund, T. Husfeldt, and M. Koivisto. Set partitioning via inclusion-exclusion. *SIAM J. Comput., Special Issue for FOCS 2006*. To appear.
- [3] H. L. Bodlaender, M. R. Fellows, and M. T. Hallett. Beyond NP-completeness for problems of bounded width: Hardness for the W hierarchy (extended abstract). In *ACM Symposium on Theory of Computing*, pages 449–458, 1994.
- [4] J. Chen, I. A. Kanj, and G. Xia. Improved parameterized upper bounds for vertex cover. In *Proc. MFCS'06*, pages 238–249, 2006.
- [5] M. Cygan and M. Pilipczuk. Faster exact bandwidth. In *Proc. WG'08*. To appear.
- [6] E. Dantsin, M. Gavrilovich, E. A. Hirsch, and B. Konev. MAX SAT approximation beyond the limits of polynomial-time approximation. *Ann. Pure Appl. Logic*, 113(1-3):81–94, 2001.
- [7] R. G. Downey and M. R. Fellows. *Parameterized Complexity*. Springer, 1999.

- [8] U. Feige. A threshold of $\ln n$ for approximating set cover. *J. ACM*, 45(4):634–652, 1998.
- [9] U. Feige. Approximating the bandwidth via volume respecting embeddings. *J. Comput. Syst. Sci.*, 60(3):510–539, 2000.
- [10] U. Feige. Coping with the NP-hardness of the graph bandwidth problem. In *Proc. SWAT’00*, pages 10–19, 2000.
- [11] U. Feige and J. Kilian. Zero knowledge and the chromatic number. *J. Comput. Syst. Sci.*, 57(2):187–199, 1998.
- [12] U. Feige and M. Singh. Improved approximation ratios for traveling salesperson tours and paths in directed graphs. In *Proc. APPROX-RANDOM’07*, pages 104–118, 2007.
- [13] F. V. Fomin, F. Grandoni, and D. Kratsch. Measure and conquer: a simple $O(2^{0.288n})$ independent set algorithm. In *Proc. SODA’06*, pages 18–25, 2006.
- [14] A. Frieze, G. Galbiati, and F. Maffioli. On the worst-case performance of some algorithms for the asymmetric traveling salesman problem. *Networks*, 12:23–39, 1982.
- [15] Y. Gurevich and S. Shelah. Expected computation time for hamiltonian path problem. *SIAM J. Comput.*, 16(3):486–502, 1987.
- [16] J. Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, 182(1):105–142, 1999.
- [17] M. Held and R. Karp. A dynamic programming approach to sequencing problems. *Journal of SIAM*, 10:196–210, 1962.
- [18] J. Kleinberg and E. Tardos. *Algorithm Design*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2005.
- [19] D. Marx. Parameterized complexity and approximation algorithms. *The Computer Journal*, 51(1):60–78, 2008.
- [20] W. Unger. The complexity of the approximation of the bandwidth problem. In *Proc. FOCS’98*, pages 82–91, 1998.
- [21] J. M. M. van Rooij and H. L. Bodlaender. Design by measure and conquer, a faster exact algorithm for dominating set. In *Proc. STACS’08*, pages 657–668, 2008.
- [22] V. Vassilevska, R. Williams, and S. L. M. Woo. Confronting hardness using a hybrid approach. In *Proc. SODA’06*, pages 1–10, 2006.
- [23] V. V. Vazirani. *Approximation Algorithms*. Springer, 2001.

A Proof of the Reduction Composition Lemma

Lemma A.1 (Reduction Composition). *If there is a (r, f) -reduction R then for any positive $k \in \mathbb{N}$ there is a (r^k, f^k) -reduction³ R' for the same problem. Moreover, if the merger of R generates a polynomial number of instances then R' has the same time complexity as R , up to a polynomial factor.*

Proof. We use induction on k . Let R and M be the reducer and the merger of R , respectively. We will describe R' , which consists of the reducer R_k and the merger M_k .

For $k = 1$ the claim is trivial. Now assume there is a (r^{k-1}, f^{k-1}) -reduction Q with reducer R_{k-1} and merger M_{k-1} .

Let I be the input instance for reducer R_k . R_k executes R_{k-1} which generates instances I_1, \dots, I_q . By induction hypothesis, $s(I_i) \leq s(I)/r^{k-1}$ for $i = 1, \dots, q$. Then for each $i = 1, \dots, q$, R_k applies R to I_i , which generates instances $I_{i,1}, \dots, I_{i,q_i}$. Note that for each i, j , $s(I_{i,j}) \leq s(I)/r^k$.

Now assume $S_{i,j}$ is a solution for $I_{i,j}$ such that $S_{i,j} \leq \alpha \text{OPT}_{I_{i,j}}$ (for minimization problem)⁴. Merger M_k applies M to every sequence of solutions $S_{i,1}, \dots, S_{i,q_i}$ and gets a resulting solution S_i for every $i = 1, \dots, q$. By the definition of reduction $m(S_i) \leq f(\alpha) \text{OPT}_{I_i}$. Then M_k applies M_{k-1} to S_1, \dots, S_q , obtaining a solution S . Then, by the induction hypothesis $m(S) \leq f^{k-1}(f(\alpha)) \text{OPT}_I$ and hence $m(S) \leq f^k(\alpha) \text{OPT}_I$, as required.

The second claim follows easily, since R_k generates q^k instances. \square

B Maximum Independent Set

Theorem B.1. *There is a polynomial-time reduction scheme for MAXIMUM INDEPENDENT SET with approximation $a(r, \alpha) = \alpha r$, for any $r \in \mathbb{Q}$, $r > 1$.*

Proof. Let $r = \frac{k}{l}$, $k, l \in \mathbb{N}$, $k \geq l > 0$. Let $G = (V, E)$ be the input graph. Reducer partitions V into k parts V_0, \dots, V_{k-1} , of size at most $\lceil |V|/k \rceil$ each. Then it creates k instances. Namely, for any $i = 0, \dots, k-1$ it creates $G_i = G[\bigcup_{j=0}^{l-1} V_{(i+j) \bmod k}]$.

Let $\text{SOL}_0, \dots, \text{SOL}_{k-1}$ be solutions (independent sets) for G_0, \dots, G_{k-1} such that $|\text{SOL}_i| \geq |\text{OPT}_{G_i}|/\alpha$. Merger simply picks the biggest solution.

We claim that for some $i = 0, \dots, k-1$, $|V(G_i) \cap \text{OPT}| \geq \frac{l}{k} |\text{OPT}|$. Indeed, otherwise all of $V(G_i)$ contain less than $l |\text{OPT}|$ copies of elements of OPT , with each element of OPT appearing in exactly l copies, hence some elements of OPT do not belong to any $V(G_i)$, a contradiction. Clearly, if $|V(G_i) \cap \text{OPT}| \geq \frac{l}{k} |\text{OPT}|$ then $|\text{OPT}_{G_i}| \geq \frac{l}{k} |\text{OPT}| = |\text{OPT}|/r$. Hence $|\text{SOL}_i| \geq |\text{OPT}|/(\alpha r)$ and the solution returned by merger will be at least that good. \square

C Coloring

There is a following simple reduction for VERTEX COLORING.

Theorem C.1. *There is a polynomial-time reduction scheme for VERTEX COLORING with approximation $a(r, \alpha) = \alpha r$, for any $r \in \mathbb{N}$, $r > 1$.*

Proof. Let $G = (V, E)$ be the input graph. Reducer partitions V into two r sets V_1, \dots, V_r , with at most $\lceil |V|/r \rceil$ vertices each and returns r instances G_1, \dots, G_r such that $G_i = G[V_i]$.

The input for merger is a coloring $c_i : V_r \rightarrow \{1, \dots, q_r\}$ for each graph G_i such that $q_i \leq \alpha \chi(G_i)$. For any i , $G_i \subseteq G$, so $\chi(G_i) \leq \chi(G)$ and further $q_i \leq \alpha \chi(G)$. Merger simply colors each $v \in V_i$ with color $\sum_{j=1}^{i-1} q_j + c_i(v)$. Clearly, it uses at most $\alpha r \chi(G)$ colors, as required. \square

³ f^k is the composition: $f^k = f \circ f^{k-1}$, $f^0 = \text{id}$.

⁴ The proof for a maximization problem is analogous

A more sophisticated reduction was found by Björklund and Husfeldt [2]. Basically, it removes maximum independent sets (found by an exact algorithm) until the graph is small enough, and the merger just colors the previously removed vertices with new colors (one per each independent set). However, taking into account current best time bounds of exact polynomial space algorithms for MAX INDEPENDENT SET and VERTEX COLORING, this reduction makes sense only when the reduction rate $r < 4.05$ (roughly). This is because for larger rates the total time of the resulting approximation is dominated by finding maximum independent sets, so we would get worse approximation guarantee with the same total time. A natural idea here is to plug our MAX INDEPENDENT SET approximation into the algorithm of Björklund and Husfeldt. By modifying their analysis we get the following theorem.

Theorem C.2. *Assume that there is a β -approximation algorithm for MAXIMUM INDEPENDENT SET with time complexity $T(n)$. Then there is an $O^*(T(n))$ -time reduction scheme for VERTEX COLORING with approximation $a(r, \alpha) = \alpha + \beta \ln r$ plus additive error 1, for any $r \in \mathbb{Q}$, $r > 1$.*

Proof. Let $n = |V(G)|$. As long as the number of vertices of G exceeds n/r reducer finds an independent set using the β -approximation algorithm and removes it from G . Let I_1, \dots, I_t be the independent sets found. Reducer returns the resulting graph G' .

Now let us upperbound t . Let $G_0 = G$ and let G_j be the graph obtained from G_{j-1} by removing I_j . Since any subgraph of G is $\chi(G)$ -colorable, for any $j = 1, \dots, k$, G_j contains an independent set of size at least $|V(G_j)|/\chi(G)$. It follows that

$$|V(G_j)| \leq \left(1 - \frac{1}{\beta\chi(G)}\right)^j n \leq e^{-\frac{j}{\beta\chi(G)}} n.$$

Hence for $j \geq \chi(G)\beta \ln r$ we have $|V(G_j)| \leq n/r$ so $t \leq \lceil \chi(G)\beta \ln r \rceil$.

Let $c : V(G') \rightarrow \{1, \dots, q\}$ be a coloring of G' , $q \leq \alpha\chi(G')$. Since G' is a subgraph of G , $\chi(G') \leq \chi(G)$ and hence $q \leq \alpha\chi(G)$ colors. Merger returns the following coloring of $V(G)$: if $v \in V'$ it has color $c(v)$ and if $v \in I_j$ it has color $q + j$. Clearly, it uses at most $\alpha\chi(G) + t \leq \alpha\chi(G) + \lceil \chi(G)\beta \ln r \rceil \leq (\alpha + \beta \ln r)\chi(G) + 1$ colors. The time complexity of the reduction is $O(nT(n))$. \square

Theorem C.3. *Given any $O^*(2^{cn})$ -time exact algorithm for MAXIMUM INDEPENDENT SET, for any $\beta \geq 1$, one can construct an $O^*(2^{cn/\beta})$ -time reduction scheme for VERTEX COLORING with approximation $a(r, \alpha) = \alpha + \beta \ln r$ plus additive error 1, for any $r \in \mathbb{Q}$, $r > 1$.*

Proof. The (β, β) -reduction from Theorem B.1 together with the $O^*(2^{cn})$ -time exact algorithm gives β -approximation for MAXIMUM INDEPENDENT SET in time $O^*(2^{cn/\beta})$. By Thm. C.2 we get the claim. \square

If we have an $O^*(2^{dn})$ -time exact algorithm for VERTEX COLORING, it makes sense to have a (a, r) -reduction of time $O^*(2^{dn/r})$. After putting $\beta = cr/d$ in Theorem C.3 (and keeping $\beta \geq 1$) we get $O^*(2^{dn/r})$ -time reduction scheme for VERTEX COLORING with approximation $a(r, \alpha) = \alpha + (cr/d) \ln r$ plus an additive error 1, for any $r \in \mathbb{Q}$, $r \geq d/c$. With the currently best known value $c = 0.288$ and $d = 1.167$ (we consider polynomial space here) it gives $a(r, \alpha) = \alpha + 0.247r \ln r$ for any $r \in \mathbb{Q}$, $r \geq d/c > 4.05$.

Note also that Theorem C.3 implies that beginning from some value of r , the simple (r, r) -reduction from Theorem C.1 outperforms the above more sophisticated one. Specifically, for the current values of c and d the threshold is $r \geq 58$.

D Reduction for Bandwidth

In this section we describe a $(2, 9\alpha)$ -reduction for BANDWIDTH. The following observation will be convenient in our proof. Observe that any ordering $f : V \rightarrow \{1, \dots, |V|\}$ corresponds to a sequence $f^{-1}(1), \dots, f^{-1}(|V|)$, which will be denoted as $s(f)$. Clearly, bandwidth of f corresponds to maximum distance in $s(f)$ between ends of an edge.

Theorem D.1. *There is a polynomial-time $(2, 9\alpha)$ -reduction for BANDWIDTH.*

Proof. Let $G = (V, E)$ be the input graph. The reducer we are going to describe creates just one instance. W.l.o.g. V does not contain isolated vertices for otherwise reducer just removes them and merger just adds them in the end of the ordering. We will also assume that $\text{bw}(G) \geq 2$ – otherwise reducer replaces G by the empty graph and merger finds the optimal solution in polynomial (linear) time.

Reducer begins with finding a maximum cardinality matching M in G . Next it defines a function $\rho : V \rightarrow V$. Note that an unmatched vertex has all neighbors matched. For each unmatched vertex v for each of its (matched) neighbors w we put $\rho(w) = v$. Consider any $uw \in M$. Assume both $\rho(u) = u$ and $\rho(w) = w$. Then u has an unmatched neighbor x and w has an unmatched neighbor y . Observe that $x = y$ for otherwise $M \setminus \{uw\} \cup \{xu, wy\}$ is a matching larger than M . In this special case redefine $\rho(u) = w$. If for some $uw \in M$ both $\rho(u)$ and $\rho(w)$ are unspecified yet we put $\rho(u) = u$ and $\rho(w) = u$ (the choice which endpoint is u and which w is arbitrary). Finally, if for some $uw \in M$ exactly one value of ρ is specified, say $\rho(u) = u$, we put $\rho(w) = u$. Now ρ is fully defined in subdomain $V(M)$. Note the following claims.

Claim 1. For any edge $uw \in M$ either $\rho(u) = u$ and $\rho(w) = u$ or $\rho(u) = w$ and $\rho(w) = w$.

Claim 2. For any $x \notin V(M)$, for any its neighbor w , $\rho(w) = w$ unless there is a triangle uwv with $uw \in M$.

Now reducer specifies the value of ρ on unmatched vertices, one by one. Let x be an unmatched vertex with $\rho(x)$ unspecified and let w be any of its neighbors. If w has another unmatched neighbor y with $\rho(y)$ unspecified we put $\rho(x) = x$ and $\rho(y) = x$. Otherwise we put $\rho(x) = w$. The way we defined ρ implies the following two claims.

Claim 3. For at least half of the vertices $\rho(v) \neq v$.

Claim 4. For any $v \in V$, $|\rho^{-1}(v)| \leq 3$.

Finally, reducer simply identifies each pair of vertices u, w (i.e. adds edges wx such that $x \in N(u) \setminus N[w]$ and removes u) such that $\rho(u) = w$. Denote the resulting graph $G' = (V', E')$. Let $V' \subset V$ in the sense that $V' = \{v \in V \mid \rho(v) = v\}$. By Claim 3 the rate of the reduction is at least 2.

Before we describe the merger let us bound $\text{bw}(G')$. Let $f : V \rightarrow \{1, \dots, |V|\}$ be the ordering of $V(G)$ with bandwidth $\text{bw}(G)$.

Claim 5. $\text{bw}(G') \leq 3\text{bw}(G) - 1$.

Proof of Claim 5. Let $g : V' \rightarrow \{1, \dots, |V'|\}$ be the ordering of V' which arranges vertices of V' in the same order as f does, i.e. $s(g)$ which is obtained from $s(f)$ by removing the vertices outside V' . We will show that g has bandwidth at most $3\text{bw}(G) - 1$. Consider any $u'v' \in E'$. Then for some $u, v \in V$, $\rho(u) = u'$, $\rho(v) = v'$ and $uv \in E$.

If both $f(u)$ and $f(v)$ are outside the interval $(\min\{f(\rho(u)), f(\rho(v))\}, \max\{f(\rho(u)), f(\rho(v))\})$, but at opposite sides of it, then $|f(\rho(u)) - f(\rho(v))| \leq \text{bw}(G)$. If both $f(u)$ and $f(v)$ are outside this interval, but at the same side of it, say at u 's side, then since $|f(v) - f(\rho(v))| \leq 2b$, we have $|f(\rho(u)) - f(\rho(v))| \leq 2\text{bw}(G)$. In any of these two cases, $|f(\rho(u)) - f(\rho(v))| \leq 2\text{bw}(G) \leq 3\text{bw}(G) - 2$. It follows that $\rho(u)$ and $\rho(v)$ are at distance at most $3\text{bw}(G) - 2$ in $s(g)$, as required.

Now assume one of $f(u)$ and $f(v)$, say $f(u)$, is in the interval $(\min\{f(\rho(u)), f(\rho(v))\}, \max\{f(\rho(u)), f(\rho(v))\})$, so in particular $u \notin V'$. Then $\rho(u)$ and $\rho(v)$ are in $s(g)$ at distance smaller by at least 1 from their distance in $s(f)$. Hence it suffices to show that $|f(\rho(u)) - f(\rho(v))| \leq 3\text{bw}(G)$.

First note that if $\rho(u), \rho(v) \notin V(M)$ then by Claim 1 $u, v \notin V(M)$ and then $M \cup \{uv\}$ is a matching larger than M , a contradiction. Hence w.l.o.g. we can assume that $\rho(u) \in V(M)$. Then $\rho(u) = u$ or $\rho(u)$ is a neighbor of u .

Now assume $\rho(v) \in V(M)$. Then also $\rho(v) = v$ or $\rho(v)$ is a neighbor of v . It follows that $|f(\rho(u)) - f(\rho(v))| \leq |f(\rho(u)) - f(u)| + |f(u) - f(v)| + |f(v) - f(\rho(v))| \leq 3\text{bw}(G)$.

Finally, let $\rho(v) \notin V(M)$. Then v is at distance at most 2 from $\rho(v)$ and hence $|f(\rho(v)) - f(v)| \leq 2\text{bw}(G)$. By Claim 2 either $\rho(u) = u$ or $v\rho(u) \in E$ so in any case $v\rho(u) \in E$ which implies $|f(v) - f(\rho(u))| \leq \text{bw}(G)$. Together we get $|f(\rho(u)) - f(\rho(v))| \leq |f(\rho(v)) - f(v)| + |f(v) - f(\rho(u))| \leq 3\text{bw}(G)$. It finishes the proof of Claim 5.

Now we describe the merger. Let $f' : V' \rightarrow \{1, \dots, |V'|\}$ be the ordering of vertices of V' with bandwidth at most $\alpha \text{bw}(G')$ for some $\alpha \geq 1$. By Claim 5, bandwidth of f is at most $\alpha(3\text{bw}(G) - 1) \leq 3\alpha \text{bw}(G) - 1$.

Merger returns the ordering f such that $s(f)$ is obtained from $s(f')$ by adding vertices of $\rho^{-1}(v) \setminus \{v\}$ (there are at most 2 of them by Claim 4) right after v . Clearly $s(f)$ is a permutation of V . Now consider any edge $uv \in E$. There are at most $3\alpha \text{bw}(G) - 2$ vertices between $\rho(u)$ and $\rho(v)$ in $s(f')$. It follows that there are at most $3(3\alpha \text{bw}(G) - 2) = 9\alpha \text{bw}(G) - 6$ vertices between $\rho(u)$ and $\rho(v)$ in $s(f)$. In other words, the distance between $\rho(u)$ and $\rho(v)$ in $s(f)$ is at most $9\alpha \text{bw}(G) - 5$. As u is at distance at most 2 from $\rho(u)$ in $s(f)$, and the same holds for v , it follows that the distance between u and v in $s(f)$ is at most $9\alpha \text{bw}(G) - 1$. It follows that f has bandwidth at most $9\alpha \text{bw}(G) - 1$. \square

The above theorem together with Reduction Composition Lemma implies:

Corollary D.2. *There is a polynomial-time reduction scheme for BANDWIDTH with approximation $a(r, \alpha) = \alpha 9^k$, for any $r = 2^k$, $k \in \mathbb{N}$.*

For any k , the above reduction gives a 9^k -approximation in time $O^*(10^{n/2^k})$ and polynomial space (using the exact algorithm of Feige and Kilian), or in time $O^*(5^{n/2^k})$ and $O^*(2^{n/2^k})$ space (using the exact algorithm of Cygan and Pilipczuk).

E $O^*(c^n)$ -time polynomial space exact algorithms for SET COVER

Clearly, to make use of universe-scaling reduction we need a $O^*(c^n)$ exact algorithm, where c is a constant. As far as we know there is no such result published. However we can follow the divide-and-conquer approach of Gurevich and Shelah [15] rediscovered recently Björklund and Husfeldt [1] and we get a $O^*(4^n m^{\log n})$ -time algorithm. If m is big we can use another, $O^*(9^n)$ -time version of it.

Theorem E.1. *There is a $O^*(\min\{4^n m^{\log n}, 9^n\})$ -time algorithm that finds a minimum-weight cover of universe of size n by a family m sets.*

Proof. The algorithm is as follows. For an instance with universe U of size n we recurse on an exponential number of instances, each with universe of size smaller than $n/2$. Namely, we choose one of m sets S and we divide the remaining elements, i.e. $U \setminus S$ into two parts, each of size at most $n/2$. We consider all choices of sets and all such partitions – there are $O(m2^n)$ of them. For each such set S and partition U_1, U_2 we find recursively \mathcal{C}_1 , an optimal cover of U_1 and \mathcal{C}_2 , an optimal cover of U_2 . Clearly $\mathcal{C}_1 \cup \mathcal{C}_2 \cup \{S\}$ forms a cover of U . We choose the best cover out of the $O^*(m2^n)$ covers obtained like this.

Consider an optimal cover OPT. To each element e of U assign a unique set S_e from OPT such that $e \in S_e$. For each $S \in \text{OPT}$ let $S^* = \{e \in U : S = S_e\}$. Let $\mathcal{P} = \{S^* : S \in \text{OPT}\}$. \mathcal{P} is a partition of U . Clearly, after removing the biggest set \hat{S} from OPT we can divide all the sets in \mathcal{P} into two groups, \mathcal{P}_1 and \mathcal{P}_2 , each covering less than $n/2$ elements from $U \setminus \hat{S}$. It follows that one of the $O^*(m2^n)$ covers found by the algorithm has weight $w(\text{OPT})$, namely the cover obtained for set \hat{S} and partition $(\bigcup \mathcal{P}_1 \setminus \hat{S}, \bigcup \mathcal{P}_2 \setminus \hat{S})$.

It is clear that the above algorithm works in time $O^*(4^n m^{\log n})$. Similarly we can also get a $O^*(9^n)$ bound – instead of at most $m2^n$ instances we recurse on at most $2 \cdot 3^n$ instances: we consider all partitions of U into three sets A, B, C . Let $|A| \geq |B| \geq |C|$. If $|A| \leq n/2$ we recurse on A, B and C and otherwise we check whether $A \in \mathcal{S}$ and if so, we recurse on B and C . \square

F Semi-Metric TSP

SEMI-METRIC TSP is a variant of the classical TRAVELING SALESMAN PROBLEM. Here we are also given n vertices and an edge weight function $w : V^2 \rightarrow \mathbb{R}$, however now the function w does not need

to be symmetric, i.e. for any x, y we may have $w(x, y) \neq w(y, x)$. Thus the instance can be viewed as a directed graph, with two oppositely oriented edges joining every pair of vertices. In this variant we assume that w satisfies triangle inequality. The goal is to find the lightest (directed) Hamiltonian cycle.

In contrast to other problems considered in this paper it is not known whether SEMI-METRIC TSP is in APX. The first approximation algorithm for this problem appeared in the work of Frieze, Galbiati and Maffioli [14] and it has approximation ratio $\log_2 n$. Currently the best result is a $\frac{2}{3} \log n$ -approximation due to Feige and Singh [12].

The best known exact algorithms are the $O^*(2^n)$ -time exponential space classical algorithm by Held and Karp [17] and a $O^*(4^n n^{\log n})$ -time polynomial space algorithm by Björklund and Husfeldt [1].

The idea of our reduction is very simple – similarly as in Section 4.1 we run a polynomial-time approximation, namely the algorithm by Frieze et al. and stop it the middle. Let us recall the algorithm of Frieze et al. It begins with finding a lightest cycle cover C_0 in G (this can be done in polynomial time by finding a minimum weight matching in a corresponding bipartite graph). Note that $w(C_0) \leq w(\text{OPT}_G)$. If the cycle cover consists of just one cycle we are done. Otherwise the algorithm selects one vertex from each cycle. Let G_1 be the subgraph of graph G induced by these vertices. Then we find a lightest cycle cover in G_1 . Note that $w(\text{OPT}_{G_1}) \leq w(\text{OPT}_G)$, which follows by the triangle inequality. Then again $w(C_1) \leq w(\text{OPT}_G)$. Again, if C_1 has just one cycle we finish, otherwise we choose a vertex from each cycle, build G_2 and so on. As the cycles in cycle covers have lengths at least 2, we finish after finding at most $\log n$ cycle covers. Finally we consider the union of all cycle covers $U = \bigcup C_i$. Clearly, U is Eulerian and we can find an Eulerian cycle E in U . Then we can transform it to a Hamiltonian cycle H by following E , but replacing paths with already visited vertices by single edges. We see that $w(H) \leq w(E)$ by triangle inequality and hence $w(H) \leq \log_2 n \cdot \text{OPT}$.

Now, assume $r = 2^k$. If we stop after creating just k cycle covers, we are left with graph G_k with $|V(G_k)| \leq |V(G)|/r$, while the cycle covers have weight at most $k \cdot \text{OPT}$. If we have an α -approximate TSP tour T in G_k , we can add it to the cycle covers and proceed as in the original algorithm. Clearly we get a Hamiltonian cycle of weight at most $(\alpha + k)\text{OPT}$.

Corollary F.1. *There is a polynomial-time reduction scheme for SEMI-METRIC TSP with approximation $a(r, \alpha) = \alpha + \log_2 r$, for any $r = 2^k$, $k \in \mathbb{N}$.*